# Rights management to enable a true Internet of Things

## Lessons from Farm to Fork and other projects

Robert Newman
Department of Computer Science, University of Wolverhampton
Wolverhampton, UK
r.newman@wlv.ac.uk

Pat Doody
Institute of Technology, Tralee
Tralee, Ireland
Pat.Doody@staff.ittralee.ie

Mira Trebar
University of Ljubljana
Ljubljana, Slovenia
Mira.Trebar@fri.uni-lj.si

Uchenna Okoke
Department of Computer Science, University of Wolverhampton
Wolverhampton, UK

***Abstract*—In this paper, we differentiate between a true 'Internet of things' and its component parts. We argue that the determining aspect of the 'Internet of Things' (IoT) is the accessibility of 'things' on the global Internet, as opposed to a simple interconnection of networked 'things'. We observe that most reported applications of the 'Internet of Things' would be more accurately described as 'Intranets of Things'. In large part, this is because the owners and operators of AIDC (Automatic identification and data capture) systems and sensor networks that in the main make up the IoT have understandable concerns about the security of their assets and therefore will limit access to that which serves their own purposes. In the wider field of the Internet 'in the large', the open mining of the Web for information has become the mainstay of many genres of research, allowing the assembly of huge corpora, enabling analytical techniques that can reveal far more information than previous limited studies. It is argued that part of the expected dividend for the IoT is to enable use on a similar scale of sensor and AIDC data, and that the results will be availability of information fusion on a huge scale, which will allow significant new knowledge to be generated. We give an example of how in one project, the RFID from Farm to Fork traceability project, this prospect has been validated to an extent on the basis that data owners voluntarily made their data available on the Web for specific purposes. Extrapolating to a more general case, we suggest that there are two services that need to be provided in order for the generalized information mining that occurs on the Internet-at-large to occur in the Internet of Things. The first is a means of cataloguing available data, which is already being addressed by services such as HyperCAT. The second is an automatic rights management service (IoT-RM), which would manage the rights and permissions and allow data owners to determine in advance to whom their data should be released, for what purposes, subject to which restrictions (such as, for instance, anonymisation) and whether any remuneration should be involved. We make some concrete proposals about the form that such an IoT-RM should take.**

***Keywords—Internet of Things, Internet of Services, rights management, security, information mining.***

## I. INTRODUCTION

The term 'Internet of Things' was originally coined by Kevin Ashton of the MIT Auto-ID Center. Writing in RFID Journal [1], he remembers first using the term in a presentation made to Proctor and Gamble. Explaining his motivation, he says:

> *Conventional diagrams of the Internet include servers and routers and so on, but they leave out the most numerous and important routers of all: people. The problem is, people have limited time, attention and accuracy—all of which means they are not very good at capturing data about things in the real world.*
>
> *And that's a big deal. We're physical, and so is our environment. Our economy, society and survival aren't based on ideas or information—they're based on things. You can't eat bits, burn them to stay warm or put them in your gas tank. Ideas and information are important, but things matter much more. Yet today's information technology is so dependent on data originated by people that our computers know more about ideas than things.*
>
> *If we had computers that knew everything there was to know about things—using data they gathered without any help from us—we would be able to track and count everything, and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling, and whether they were fresh or past their best.*
>
> *We need to empower computers with their own means of gathering information, so they can see, hear and smell the world for themselves, in all its random glory.*

It's clear that the original conception had at its core the Internet, and furthermore, some developing idea that if this rich

ecology of ‘things’ could be interconnected, great things would happen. This presentation took place in 1999. Since then, 16 years have passed and while the technology of ‘things’ has progressed a great deal, in terms of the available processing resource in devices such as sensor nodes and AIDC systems, as well as ever more ingenious sensing devices, the Internet of Things has not progressed to the extent where computers can ‘see, hear and smell the world for themselves’.

To a large extent, this is not a problem of technology. As observed above, the computational power of sensor nodes and embedded systems in general increases year on year. Further, the Internet is by far the cheapest method of interconnection for anyone operating a network of sensors and other computational resources, which has resulted in many such networks being connected to the ultimate user of the data by means of the Internet. Thus these devices are already ubiquitous on the internet, however rather than making up an ‘Internet of Things’ they are configured as a collection of ‘Intranets of Things’, with security measures in place expressly to ensure that the data they provide is not openly and generally available.

The purpose of this paper is to consider the organizational blocks on the achievement of a true ‘Internet of Things’ as originally conceived by Ashton and then to propose an additional part of Internet of Things architectures, a component concerned with automatic rights management, which we argue will provide the technological underpinnings required to allow removal of these blocks. We start by reviewing some of the potential benefits of information mining in the Internet of Things. Subsequently we review the commonly expressed security concerns and the technological solutions proposed to address them. A specific example of the operation of collaborative data sharing and limited consensual data mining is examined. The selected example is the RFID from Farm to Fork (RFID-F2F) project conducted by a team including two of the present authors. The operation of this system depended on information mining, but the information providers were knowing participants and had provided this information on the open Internet just so that it could be mined. To date, much activity concerned with enabling ‘thing’ owners to make their data available on the Internet has been directed at solving the problems of discoverability, particularly with the provision of cataloguing services such as HyperCAT [2], so in order to provide context for the requirements for rights management, these systems are discussed. From this and consideration of the RFID-F2F information owners concerns, and the project activity required to address them we develop a set of requirements for automatic rights management service on the Internet of Things (IoT-RM).

## II. Information Mining the Internet of Things

The idea of information mining was central to Ashton’s original vision. Since then, many reports have reiterated this theme, or simply built it implicitly into the justification for the work. Such implicit statements are often simply projections of the future value of the Internet of Things as a massive connection of devices on the Internet without any explicit statement on how that interconnection will produce beneficial results. A rich seam of such statements is provided by the report of the Cluster of Research Projects on the Internet of Things [3]. For instance, Santucci quotes Haladjian [4]:

> *And men got dreaming. Shouldn't there be a network that made all my devices collaborate at all times, converse spontaneously among themselves and with the rest of the world, and all together make up a kind of single virtual computer – the sum of their respective intelligence, knowledge and know how?*

While information mining is not explicitly mentioned, the notion that a computing device might ‘spontaneously’ reach out for information from ‘the rest of the world’ to increase ‘intelligence, knowledge and know how’ seems very much to be what information mining is about.

Santucci notes that:

> *One of the main challenges of the Internet of Things is therefore to transform connected objects into real actors of the Internet by developing and implementing appropriate applicative design methodologies This shift of paradigm involves major societal and ethical challenges that loom ahead and need to be tackled, certainly at European level but also at global level. The metamorphosis of objects, if left without any regulation or interference, might give rise to a genuine, extensive surveillance society.*

Again, what is clearly implied is the idea that the data from devices will be freely accessible on the Internet, and used sufficiently widely that this use could be a significant threat to personal liberties.

It seems likely that information mining on the Internet of Things will be a rich source of knowledge. It is not an outrageous speculation to suggest that a very rich image of world trade could be gained by mining information from EPC (Electronic Product Code) tags or that data from the temperature and air pressure sensors on the Internet could be mined to produce a very detailed picture of microclimatic conditions. The techniques of information mining can allow very powerful derivation from many kinds of data. This is the context within which Barnaghi et. al. discussed the need to develop a semantic layer for the Internet of Things [5]. They note that:

> *... the current initiative on building the IoT (or more general, the future Internet) demands application and service platforms which can capture, communicate, store, access and share data from the physical world. This will create new opportunities in a long list of domains such as e-health, retail, green energy, manufacturing, smart cities/houses and also personalized end-user applications.*

There are already some examples reported of information mining type activity over domains that could be included within the heading ‘Internet of Things’. Datcu et. al. report information mining in remote sensing archives. The field of ‘remote sensing’ somewhat predates the ‘Internet of Things’, but is in some ways a model for how information synthesis from devices on the Internet might operate. In this case, the devices are various cameras, resulting in databases of images.

In this work, a three stage method is reported. Primitive image features and meta-features are extracted, by clustering

these features an image vocabulary is extracted and finally Bayesian techniques are used to attach user-defined semantics to these vocabulary terms. By this means, the system learns via an interactive process the means to search very large databases for features deemed to have some semantic inference by the human user. It is easy to see how similar techniques could be employed for more general sensor data (if only because humans will often interpret data from such sensors by graphing) potentially resulting in identification of trends and patterns occurring in many world-wide phenomena, in fact anything that could be tracked using the devices associated with the Internet of Things. In the next section, we discuss a more specific case of an application designed around information mining on the Internet of Things.

## III. RFID FROM FARM TO FORK – AN EXAMPLE OF COLLABORATIVE INFORMATION MINING

The RFID from Farm to Fork project is discussed by Cuinas et. al. [6]. The aim of the project was to allow consumer visibility of the production history and handling conditions in the supply chain. In order to allow this to happen, food manufacturers tracked their production process using a combination of AIDC and sensing technology. The AIDC technology was based around 1-D and 2-D barcodes, QR (Quick Response) codes and RFID or a mixture, according to the specific requirements in the manufacturing context. The sensing technology ranged from simple temperature sensors in the cold chain to wireless sensor networks monitoring growing conditions in vineyards. All of this information was formatted according to GS1 EPCIS (Electronic Product Code Information Services) concepts [7], in most cases using the open source Fosstrack implementation of EPCIS Repository as well as Query and Capture clients. The EPCIS standard assigns uniform resource identifiers (URIs) to individual assets, which is a globally unique name including a uniform resource locator, which can be used to locate the 'location' of the information on the World Wide Web.

A consumer on purchasing a product scans a QR code~~,~~ Datamatrix or NFC tag, which provides the URI of the product just bought. This is a serialized URI, which may identify the specific product, delivery or production batch, depending on the traceability requirements of the product manufacturer. As often as not, this depends on the value of the item concerned. The query proceeds as follows. First the URL in the product URI is used to access the EPCIS repository at that location, which returns the information available for that product, which might include production dates, temperature logs, other sensor data and URIs for the ingredients. In turn, the product databases corresponding to those URIs are accessed and the product history of the ingredients retrieved. Proceeding in this way, the traceability process in the supply chain provides the entire history of the product and all of its ingredients can be discovered, hence the project title 'from Farm to Fork'. The user interface to this information is a custom website for that product, which is generated in real time by an application called 'Identity Explorer', and thus is usable on any device (smart phone) which can run a web browser. If the device also includes a camera (to scan QR code or read a Datamatrix) or NFC reader, then it is able to display all the available product information on that product. Availability of that information depends on the willingness of its owner to make it available on the open Internet, where the Identity Explorer can find it. There are many reasons why commercial concerns, such as food manufacturers, might wish to keep their production information confidential. These include maintenance of trade secrets, a concern that data mining of this information might reveal information on their business that they do not want to release and a feeling that revealing information openly poses some kind of unidentified threat.

To counteract these concerns, it was necessary to provide the participants with some real business advantage that accrued from the release of the information. In the case of the food industry, which has been affected by a number of scares concerning the authenticity and origin of foodstuffs, that advantage was increased marketability of their product. The presentation of full traceability information to the consumer provides a large increase in consumer confidence and increases the saleability of the product. This was particularly pronounced for manufacturers of premium products, since it provided a means of clearly demonstrating that they were using superior ingredients and particular manufacturing processes.

Having provided a business case for exposing the data, it was necessary to provide reassurance that only the data that the manufacturers wanted released was accessed in the EPCIS. The information had been sourced in a variety of ways. Some of the larger manufacturers involved already ran processes with full traceability and extensive use of sensor equipment in the manufacturing process. In this case, all that was necessary ~~was~~ to select the data required for consumer traceability and perform a translation to the required EPCIS formats if necessary. With some smaller manufacturers traceability and process tracking was done entirely with paper based systems. While it would have been possible to transpose the data from those systems to an EPCIS server, the practice was to provide automatic tracing and sensing by installing some economical AIDC and sensing systems, which interfaced directly to the EPCIS system.

We argue that the RFID from Farm to Fork system was part of the Internet of Things, in that AIDC and sensor data was available directly on the Internet and that it was an example of information mining, in that it involved the real-time assembly of information from many services, with that information gathering being performed automatically. However, as an information mining application it was atypical, relying on data from organisations which had agreed in advance to make a subset of their production data available for the specific purpose of providing traceability, working with the system designers to determine which data was to be revealed and uploading that data to specialized servers, rather than allowing their working databases to be mined.

This contrasts with more usual instances of data mining, where those performing the mining decide what they will do with the data, and devise strategies using search engines to trawl the Internet for suitable data. If there were not data owners making their datasets available on the open Web, such strategies simply would not work. We suggest therefore that if the value of information mining the Internet of Things is to be realised that the owners of the data to be mined must have a

way to gain the same kind of confidence in the use of their data that those taking part in the RFID-F2F project did. This can be summarized as follows:

- They should be able to gain some benefit from the use of their data.
- They should have some control over who uses their data and for what purpose.
- It should be possible to maintain control of the data, for instance in some cases it may be necessary to anonymise data.

To this end, we propose that Internet of Things architectures need to be expanded to include a layer that specifically deals with these concerns, a digital rights management layer implemented as part of the Internet of Services that connects things and people.

## IV. Security Concerns

Security applied to the Internet of Things is an active strand of research. In his study of research directions for the Internet of Things, Stankovich identifies security as a major concern [8]. The IoT was identified as being vulnerable due the physical accessibility to sensors, actuators and objects, and the openness of the system. However, the physical accessibility is a particular concern of 'things' in general, and not a matter of their connection to the Internet. Roman et. al, provide an overview of security concerns in the Internet of Things, covering concerns including classical data security, but also the issues of privacy and identity [9]. They note that:

> *The IoT's highly distributed nature and use of fragile technologies, such as limited-function embedded devices in public areas, create weak links that malicious entities can exploit. Easily accessible objects in unprotected zones, such as city streets, are vulnerable to physical harm. Like compromising botnets, some objects would try to hinder services from the inside. Additional threats include the existence of a domino effect between intertwined services and user profiling through data collection and other methods.*

Discussing the matter of privacy, they identify the potential for attacks on the IoT to yield personal information:

> *Privacy is one of the most sensitive subjects in any discussion of IoT protection. The data availability explosion has created Big Brother-like entities that profile and track users without their consent. The IoT's anywhere, anything, anytime nature could easily turn such practices into a dystopia. Users would have access to an unprecedented number of personalized services, all of which would generate considerable data, and the environment itself would be able to acquire information about users automatically.*

The question of identity and ownership of data was a key concern in the RFID-F2F project, but is rarely considered in great detail in the literature of IoT security and privacy. Roman, et. al., do however put some thought into this, particularly introducing the notion of 'identity shadowing', in which 'a user projects his virtual identity onto logical nodes'. From the point of view of this study, we would interpret that as proposing that the owner of the nodes is the owner of the data, and takes the rights and responsibilities for that data, which flow from that ownership. It is not legally a formal copyright situation, but may be to some extent covered under EU database rights [10].

One way of classifying attacks is by the method of attack Kopetz classified various security attacks that can be as [11]:

- Malicious attack: Where an adversary inserts malicious code.
- Spoofing attack: the adversary masquerade as a legitimate user in order to gain unauthorized access to a system.
- Password attack: The password of the system is been guessed. There are two versions of this kind of attack. Dictionary attacks and brute force attacks.
- Cipher-Text attacks: The attacker assumes to have access to the cipher text and tries to deduce the plain text and possibly the encryption key from the cipher text.
- Denial of Service attacks: The attack tries to make a computer system unavailable to its users by jamming the network.
- Botnet attack; a set of infected networked nodes like thousands of PC's that are under the control of an attacker.

Given the exposure that has been given to this kind of concern, it is not surprising that many owners of device networks are concerned about the protection of their security. Taking this into account, in addition to the commercial concerns discussed above, it is understandable that there is considerable reluctance to make the data for networks of devices available on the open Internet. We propose that standard architectures for the IoT should include a service layer for rights management to be placed between network and application layer.

## V. Automatic Rights Management for Things

The deliverables of the EU Framework 7 project 'Internet of Things – Architecture' [12] produced a complete architectural framework for the Internet of Things. One of the major concerns addressed, given its own separate deliverable was 'Privacy and Security' [13]. In the executive summary, the importance of these matters is stressed:

> *Security is an important cornerstone for the Internet of Things (IoT). More specific, all common aspects of security must be regarded. With the huge amount of data created by IoTs, integrity of data and trust in the services offering the data is crucial. Further, to protect important data and user interests, confidentiality of data and privacy of users must be ensured. In addition to integrity and confidentiality, each request and response inside the*
>
> *IoT has to be authenticated in a proper and secure way.*

The stress is on stopping unauthorised access, rather than finding means of providing the widest authorised access compatible with the need to address security and privacy.

In their motivation for the need for a semantic layer within the Internet of Things, Barnaghi et. al. [5] discuss the various functions that will be required to enable widespread use of information mining type functions on the Internet of Things, including semantics for interoperability, IoT data integration, IoT data abstraction and access, Resource/service search and discovery. This is beginning to acknowledge that discoverability is a major concern for an effective use of IoT. Consequently, in recent times there has been some interest in cataloguing services, such as HyperCAT [14], which provide for discovering IoT resources. HyperCAT provides a means for resource owners to make their resources discoverable by listing them in an open catalogue. However, it does not address the motivation for having them discoverable. On the basis of our experience with RFID-F2F, we believe that resource owners will require the following reassurances:

1. That they will receive some benefit for making their resources discoverable.
2. That they will be able to separate data that is discoverable from that that isn't and provide secure protection for the latter.
3. That they may wish to restrict the community by whom this information is discoverable.

On the basis of the concerns detailed in the IoT-A deliverables, particularly to do with privacy:

4. That it must be possible when required to process the data in order to anonymise it, that is to ensure that data processing cannot reveal confidential data on individuals or organisations from this data.

Also of importance is the manner of operation of such an architectural component. Information mining would lose much of its power if it was necessary for each transaction in the search to be individually negotiated with the data owner. For this reason, the decisions on whether or not to grant access to a search need to be performed automatically which in turn necessitates that the data owner has predetermined the access criteria. The rights management system then becomes a semantic issue, classifying the enquirer into one of the permitted classes, and making decisions on which rights to grant and whether there needs to be any remuneration.

We foresee the operation of the rights management layer as follows.

### *A. Benefit*

In the field of open source software and 'creative commons' media, there is frequently a layered model of access rights. That is, free access is offered to some classes of user (very often non-commercial), with different terms, including payment, required for commercial use. We would envisage that the IoT-RM solution for IoT data would need to follow a similar layered approach, most probably and most easily based on some standardised licence models similar to the 'Creative Commons' licences. In the case for which remuneration was required, there would need to be a mechanism for funds transfer built into the search mechanism. We could also speculate that, similar to the wider Internet, there may be alternative models for 'monetisation', including advertising and affiliation.

### *B. Security of undiscoverable data.*

As discussed above, security remains a prime concern for thing owners, however speculative the security threats. In a scenario where access decisions concerning all or some of the data are being made automatically, there needs to be confidence that the data and systems to which access is not being granted remain secure. Given that much of the security measures proposed involve encryption of some kind, we could suggest that the solution to this problem might be selective or layered encryption, that is, that the data be encrypted in a layered manner (also known as multiple encryption) with different keys associated with different kinds of access permission. This is an approach already used in automatic image rights management [16] and is core to the 'Onion Router' [17].

### *C. Identifying the searcher's community*

The question of identifying searchers to whom access of all or part of the data should be allowed maybe much more complex than is common in simple media rights management. For example, many sensor webs will fall into protected categories under various states security legislation, such as the USA's National Security Laws [18]. This type of resource will yield a great deal of useful information to many scientific surveys, so a great deal of the value of the IoT would be lost if they were denied under the proposed mechanism to those who were authorised to use them. Generally the legal principle is that users of the data must request and be authorised in advance, but it is in the nature of automated web searches that there is no 'in advance' so far as the human operator is concerned. Thus, it must be possible for the identity and authority of the searcher to be determined reliably and quickly within the context of an automatic protocol. It can be imagined that national security agencies would demand a high degree of reliability from such a service, and this this would seem to be an area requiring ongoing research.

### *D. Anonymisation*

The requirement for anonymisation really stems from the requirement to respect privacy. Often, it is proposed that AIDC data will constitute a direct privacy threat, though in truth it is only that AIDC data which directly identifies individuals that does so. More problematic is the possibility of inferred identity, relating from data fusion across different data sets, which might come from different data owners. The problem with ruling out access to some types of data on the basis of privacy is that this also negates some of the power of web information mining. For example, it is not hard to see than some very large and detailed epidemiological studies could be made by web searches of such things as environmental and pollution sensing networks and fusion of that data with health records. If too simplistic a model of privacy protection is taken than such studies would become difficult or impossible. Therefore, we suggest that necessary research should begin to develop a

layered model for privacy protection of individuals that will allow the release of information, properly anonymised, to allow this type of search. The model would need to take into account the possibility of inferred identity as well as direct revelation of identity.

## VI. Conclusion

In this paper we have argued the need for IoT architectures to include a layer of automatic rights management implemented as IoT-RM and that the full potential of the Internet of Things, as envisaged by Ashton and other, will not be fulfilled without this. We have also found that the need for this component has not featured in previous studies of IoT architectures or proposals for them, including that from the IoT-A project, Framework 7's major effort at IoT model normalisation. We therefore propose that the IoT community include and adopt this as an important line of research in the coming period when organisations and industry are concerned.

## Rererences


[1] Ashton, K. That 'internet of things' thing. *RFiD Journal*, *22*(7), pp. 97-114, June 2009.

[2] Blackstock, Michael, and Rodger Lea. "IoT interoperability: A hub-based approach." *Internet of Things (IOT), 2014 International Conference on the*. IEEE, 2014.

[3] Sundmaeker, H, Guillemin, P, Friess, P, Woefflé, S.*Vision and challenges for realising the Internet of Things*. Vol. 20, no. 10. EUR-OP, 2010.

[4] Santucci, G. The Internet of Things: Between the Revolution of the Internet and the Metamorphosis of Objects in Sundmaeker, H, Guillemin, P, Friess, P, Woefflé, S.*Vision and challenges for realising the Internet of Things*. Vol. 20, no. 10. EUR-OP, 2010.

[5] Barnaghi, Payam, Wei Wang, Cory Henson, and Kerry Taylor. "Semantics for the Internet of Things: early progress and back to the future." *International Journal on Semantic Web and Information Systems (IJSWIS)* 8, (1) , pp. 1-21, 2012.

[6] Cuinas, Inigo, Robert Newman, Mira Trebar, Luca Catarinucci, and Alejandro A. Melcon. "Rfid-based traceability along the food-production chain [Wireless Corner]." *Antennas and Propagation Magazine, IEEE* 56(2), pp. 196-207. 2014.

[7] GS1. EPC Information Services (EPCIS) Version 1.1 Specification 2, GS1 AISBL, Brussels. Dec 2013.

[8] Stankovic, J.A. Research directions for the internet of things. Internet of Things Journal, IEEE, 1(1), pp. 3-9. 2014

[9] R. Roman, P. Najera, and J. Lopez, "Securing the Internet of Things", IEEE Computer, vol. 44, pp. 51 -58, 2011.

[10] European Commission, "Protection of Databases", http://ec.europa.eu/internal_market/copyright/prot-databases/index_en.htm, October, 2014.

[11] Kopetz, H., Real-time systems: design principles for distributed embedded applications, Springer Science & Business Media, 2011.

[12] IoT-A Project, Final Architectural Reference Model for the IoT. Deliverable 1.5 Available at http://www.iot-a.eu/public/public-documents/d1.5/view, 2012.

[13] IoT-A Project, Concepts and Solutions for Privacy and Security in the Resolution Infrastructure. Deliverable 4.2 Available at http://www.iot-a.eu/public/public-documents/d4.2/view, 2012.

[14] Lea, R., *HyperCat:an IoT interoperability specification.* IoT ecosystem demonstrator interoperability working group. Available at http://eprints.lancs.ac.uk/69124/, 2013.

[15] Craetive Commons, "About the Licenses", https://creativecommons.org/licenses/.

[16] Tosun, A. S. "Efficient multi-layer coding and encryption of MPEG video streams." In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, vol. 1, pp. 119-122. IEEE, 2000.

[17] Bauer, K. S., Sherr, M., Grunwald, D. "ExperimenTor: A Testbed for Safe and Realistic Tor Experimentation." *CSET*. 2011.

[18] Dycus, Stephen. *National security law*. Aspen Law & Business, 2007.